

doi:10.3969/j.issn.1671-3168.2023.06.002

基于 Landsat 的香格里拉市高山松地上生物量动态研究

罗利彬¹, 张加龙²

(1. 迪庆州林草种苗和国有林场管理工作站, 云南 迪庆 674499; 2. 西南林业大学, 云南 昆明 650224)

摘要:采用不同时期的 Landsat5 TM 卫星遥感影像数据和 1987—2007 年 5 期云南省香格里拉市森林资源连续清查样地调查数据,通过数据筛选,应用随机森林算法(RF)、梯度提升回归树算法(GBRT)等相关性分析模型,估测 1987—2007 年间云南省香格里拉市高山松地上生物量动态变化规律。结果表明,GBRT 算法的估测模型效果最好,决定系数 R^2 为 0.99,预估精度 P 为 70.07%;RF 算法次之,决定系数 R^2 为 0.89,预估精度 P 为 66.10%。1987—2007 年的 20 年间,香格里拉高山松地上生物量总量经历了先减又增的过程,1987、1992、1997、2002、2007 年地上生物量分别为 1 023.29、1 022.38、1 011.73、1 018.02、1 019.33 万 t。但截至 2007 年,高山松地上生物量仍然未恢复到 1987 年水平。结合 20 年的林业发展过程,对高山松地上生物量动态变化原因进行简要分析,对后续研究提出了建议。

关键词:高山松;地上生物量;动态变化;梯度提升回归树算法(GBRT);香格里拉市

中图分类号:S791.2594;S718.556;S771.8 文献标识码:A 文章编号:1671-3168(2023)06-0007-06

引文格式:罗利彬,张加龙. 基于 Landsat 的香格里拉市高山松地上生物量动态研究[J]. 林业调查规划,2023,48(6):7-12. doi:10.3969/j.issn.1671-3168.2023.06.002

LUO Libin, ZHANG Jialong. Aboveground Biomass Dynamics of *Pinus densata* in Shangri-La Based on Landsat[J]. Forest Inventory and Planning, 2023, 48(6):7-12. doi:10.3969/j.issn.1671-3168.2023.06.002

Aboveground Biomass Dynamics of *Pinus densata* in Shangri-La Based on Landsat

LUO Libin¹, ZHANG Jialong²

(1. Deqen Administration of Forest and Grass Seedling and State-owned Forest Farm, Deqen, Yunnan 674499, China;
2. Southwest Forestry University, Kunming 650224, China)

Abstract: Based on Landsat5 TM satellite remote sensing image data from different periods and forest resource inventory data of five periods from 1987 to 2007 in Shangri-La City, Yunnan Province, the dynamic change law of aboveground biomass of *Pinus densata* in Shangri-La was estimated by using correlation analysis models such as random forest algorithm (RF) and gradient boost regression tree algorithm (GBRT). The results showed that the estimation model of GBRT algorithm had the best effect, with the determination coefficient R^2 of 0.99 and the estimation accuracy P of 70.07%; secondly, RF algorithm, the determination coefficient R^2 was 0.89, and the prediction accuracy P was 66.10%. In the past twenty years, the total aboveground biomass of *Pinus densata* in Shangri-La experienced a process of first decreasing and then increasing, with 10.232 9 million tons, 10.223 8 million tons, 10.117 3 million tons,

收稿日期:2022-06-23;修回日期:2022-12-27.

第一作者:罗利彬(1987-),男,云南迪庆人,硕士,工程师.从事森林保护、林草种苗管理工作. Email:447332433@qq.com

责任作者:张加龙(1981-),男,云南昆明人,博士,教授.主要研究方向为基于遥感的森林生物量建模、土地利用与覆盖变化、森林覆盖分类、森林资源信息管理、时空 GIS 等. Email:38398764@qq.com

10. 180 2 million tons and 10. 193 3 million tons in 1987, 1992, 1997, 2002 and 2007 respectively. But by 2007, the aboveground biomass of *Pinus densata* had not yet recovered to the level of 1987. Based on the 20-year forestry development process, this paper analyzed the reasons for the dynamic changes in aboveground biomass of *Pinus densata*, and put forward suggestions for further research.

Key words: *Pinus densata*; aboveground biomass; dynamic changes; gradient boost regression tree algorithm; Shangri-La City

森林生物量是衡量森林生态系统物质积累与物质循环的重要指标。探索森林生物量的动态变化对于研究森林碳循环具有重要意义。森林生物量是森林生态系统在长期的生产和代谢过程中积累的能量物质,是森林固碳能力的集中体现,通常以单位面积森林群落落在一定时间内积累的有机质总量来表示,单位为吨每公顷(t/hm^2)^[1]。森林生物量的准确估测对研究森林生态系统碳循环和区域性气候变化具有重要意义^[2-3]。云南香格里拉市作为滇西北低纬度高海拔温带、寒温带森林典型分布区,长期以来对其森林生物量的研究较少。2016年,陆驰等^[4]建立了基于遥感因子的高山松生物量估测的非线性和线性模型,采用“交叉验证法”评价模型的预测能力,并认为线性模型的预测效果比非线性模型好。2013年,孙雪莲等^[5]以云南省香格里拉市高山松林为研究对象,以地面样地调查和单木生物量实测数据为基础,充分考虑遥感变量,运用参数和非参数模型构建高山松林生物量遥感估测模型,并反演香格里拉市高山松林生物量,实现森林生物量从小尺度向大区域的转变。徐云栋等^[6]以 Landsat TM (30 m) 影像和 DEM (30 m) 数据为信息源,结合森林资源二类调查数据和地面样地实测数据,借助 MATLAB 平台,在前期进行基于遗传算法(genetic algorithm, GA)和粒子群算法(partical swarm optimization, PSO)优化 BP(back propagation)神经网络模型基础上,采用决定系数(coefficient of determination, R^2)、均方根误差(root mean square error, RMSE)及预测精度(P)3个指标对优化后的 BP 神经网络模型进行评价,并建立了研究区高山松蓄积量估测模型。作者认为,遗传算法优化下的 BP 神经网络估测香格里拉市高山松蓄积量精度优于粒子群算法优化下的 BP 神经网络估测精度。张加龙等^[7]2019年使用 Landsat OLI 数据,以香格里拉市高山松为研究对象,采用对比参数和非参数的建模方法,得出 GBRT 为理想生物量估测模型。

本文在前人研究的基础上,使用不同时期的 Landsat5 TM 卫星遥感影像数据和 1987—2007 年间 5 期云南省香格里拉市森林资源连续清查样地调查

数据,通过数据筛选,应用随机森林算法(RF)、梯度提升回归树算法(GBRT)等相关性分析模型,估测 1987—2007 年间云南省香格里拉市高山松地上生物量动态变化规律,以期能为该区域下一步制定森林经营方案及保护规划提供参考。

1 研究区概况

香格里拉市隶属云南省迪庆藏族自治州,地理坐标为东经 $99^{\circ}20'$ ~ $100^{\circ}19'$,北纬 $26^{\circ}52'$ ~ $28^{\circ}52'$,总面积 $11\,613\,km^2$ ^[8]。位于横断山脉南麓,滇西北三江并流腹地,东与四川省稻城、木里二县接壤,西、南与丽江市、维西县接壤,北与德钦县接壤。根据 2020 年香格里拉市森林资源主要指标统计数据公报,全市土地面积 $114.17\,万hm^2$,森林面积 $89.75\,万hm^2$,蓄积量 $1.39\,亿m^3$,有林地面积 $96.15\,万hm^2$,特殊灌木林面积 $12.39\,万hm^2$,森林覆盖率 74.99% 。

2 数据来源及预处理

2.1 固定样地数据

固定样地数据来源于 1987—2007 年每 5 年一次的香格里拉市森林资源连续清查(national forest inventory, NFI)数据,包含了连续清查样地、样木、成果分析、统计报表、遥感和基础地理数据等信息^[10]。该固定样地是从云南省范围内按 $6\,km \times 8\,km$ 均匀布设的边长 $28.28\,m$ 、面积 $0.08\,hm^2$ 的正方形固定样地中选取的^[11]。高山松地上生物量计算采用 2016 年孙雪莲等^[5]建立的高山松单木生物量计算方程转换。计算公式为:

$$W = 0.073 \times DBH^{1.739} \times H^{0.880} \quad (1)$$

式中: W 为单木生物量(kg); DBH 为胸径(cm); H 为树高(m)。

利用上述方程式计算出高山松单木生物量,最后将样地内所有高山松的单木生物量进行汇总,得到标准样地尺度内的高山松地上生物量总量。

2.2 遥感影像数据

遥感影像数据来源于美国地质勘探局(United

States Geological Survey,简称 USGS)发布的 1987 年、1992 年、1997 年、2002 年和 2007 年 Landsat5 TM 数据。经过几何矫正配准,筛选云量小于 15% 的影像,数据获取年份如表 1 所示。

表 1 获取的研究区 Landsat TM 影像数据
Tab. 1 Collected images of Landsat TM in the study area

年份	ID	获取时间	云量/%
2007	LT51310412007060BJC00	2007-03-01	1.00
	LT51320402007003BJC01	2007-01-03	23.00
	LT51320412006288BJC00	2006-10-15	15.00
2002	LT51310412002302BJC00	2002-10-29	0.13
	LT51320402002005BJC00	2002-01-05	0.15
	LT51320412002005BJC00	2002-01-05	2.80
1997	LT51310411997320BKT01	1997-11-16	7.00
	LT51320401997279BKT00	1997-10-06	16.00
	LT51320411997311BKT00	1997-11-07	4.00
1992	LT51310411991320BKT00	1991-11-16	8.05
	LT51320401991311BKT00	1991-11-07	2.44
	LT51320411991311BKT00	1991-11-07	5.24
1987	LT51320411987364BKT00	1987-12-30	10.60
	LT51320401987364BKT00	1987-12-30	23.89

2.3 数字高程(DEM)数据

数字高程数据计算平台为谷歌地球引擎,数据源为 NASA 喷气推进实验室(jet propulsion laboratory, JPL)航天飞机雷达地形测绘使命(shuttle radar to-

pography mission, SRTM)项目提供的 SRTM Version 3 产品(SRTM Plus),分辨率为 30 m,该数据集使用开放的源数据进行了空隙填充。香格里拉市不同年分高山松分布如图 1 所示。

2.4 数据预处理

对于 Landsat 数据源的预处理过程包括辐射定标、大气校正、几何校正。由于研究区地形起伏较大,需要通过地形校正来实现消除地形对遥感影像的影响,参考已经校正好的 SPOT-5 影像,将所有影像校正到北京 1954 坐标系,每景图像至少选择 30 个以上的地面控制点,使用双线性内插法将影像重采样为 30 m×30 m,误差控制在一个像元内,采用二项式方法来校正图像^[12]。

3 研究方法

通过构建随机森林模型(RF)、梯度提升回归树模型(GBRT),将模型的拟合情况进行对比分析,以精度较高的模型进行 5 期高山松生物量估测。

3.1 生物量模型的建立

3.1.1 基于随机森林模型 RF 算法的模型构建

1) 特征变量筛选

在本究中,利用随机森林算法构建高山松地上生物量估测模型时,使用变量重要性评分(variable importance measures, VIM)方式对输入的 126 个特征变量的训练数据进行筛选。在进行第一次分析时,输入模型的变量为 126 个,当变量的累积贡献度达

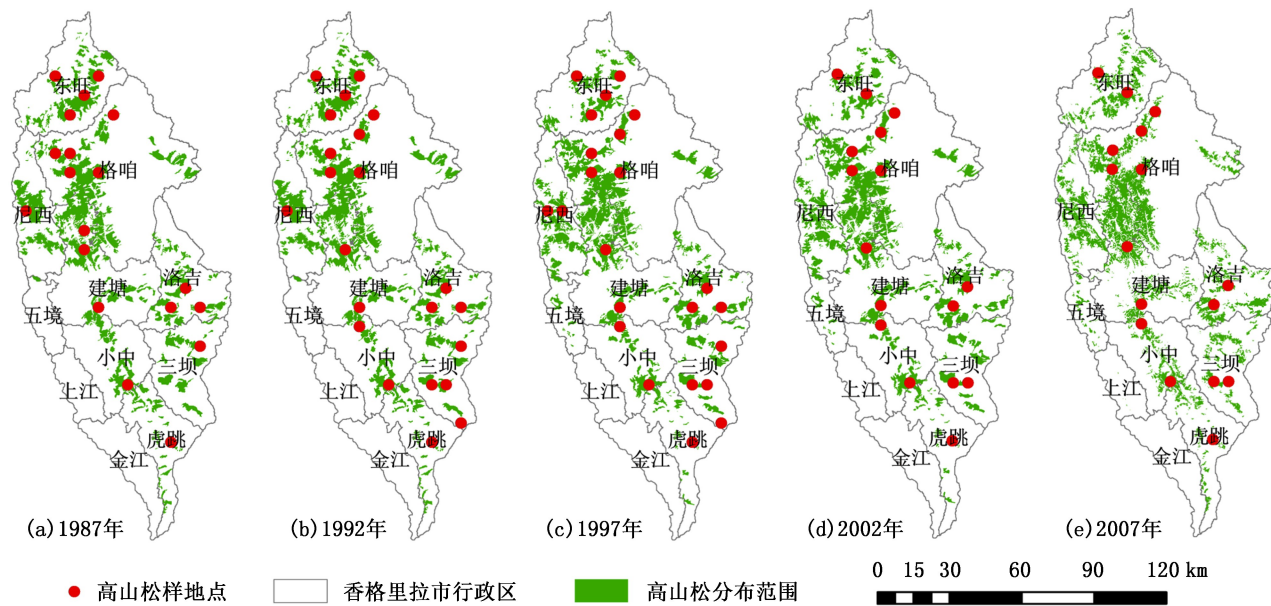


图 1 不同年份的高山松分布
Fig. 1 Distribution of Pinus densata in different years

到 80% 时,共有 38 个特征参与重要性评分积累。继续使用这前 38 个变量进行第二次模型构建,当变量的累积贡献度达到 80% 时,共有 17 个特征参与重要性评分积累。继续使用这前 17 个变量进行第三次模型构建,当变量的累积贡献度达到 80% 时,共有 10 个特征参与重要性评分积累。进一步使用这前 10 个变量继续进行第四次模型构建。

2) 模型拟合

可以看出,当模型进行第三次运算,参与运算的变量数为 17 时,模型具有最高的预测精度和最低的相对均方根误差。从拟合优度和估测均方根误差的评价角度来看,第三次运行结果略低于第四次运算结果,但因二者差异较小,本研究将选择第三次运算结果进行高山松地上生物量估测模型的构建(表 2)。

表 2 随机森林算法模型多次构建效果

Tab. 2 Multiple construction effects of random forest algorithm model

运算次数	变量数	R^2	RMSE /($t \cdot \text{hm}^{-2}$)	P /%	rRMSE /%
第一次运算	126	0.86	36.48	57.80	49.52
第二次运算	38	0.87	35.79	61.90	49.22
第三次运算	17	0.89	32.26	66.10	40.92
第四次运算	10	0.99	29.72	64.60	44.77

此时参与模型构建的 17 个特征变量为: Elevation、Aspect、TM73、R9T3VA、R9T2CC、R9T4CC、R5T3CC、R5T5CC、R5T4VA、TM437、R5T7CC、ND32、R5T4HO、R5T7CO、R9T1VA、R9T7ME、R9T1EN。在模型中变量的贡献度排序如图 2 所示。

构成模型各特征变量中,R9 和 R5 分别代表灰度共生矩阵计算时大小为 9×9 和 5×5 的窗口;TM1、TM2、TM3、TM4、TM5、TM7 分别代表 Landsat5 TM 影像的第 1、2、3、4、5、7 波段,即 Blue 波段、Green 波段、Red 波段、NIR 波段、SWIR1 波段和 SWIR2 波段;CC 代表纹理特征相关性,CO 代表纹理特征对比度,DI 代表纹理特征相异性,EN 代表纹理特征熵,HO 代表纹理特征均匀性,ME 代表纹理特征均值。

3.1.2 基于梯度提升回归树模型 GBRT 算法的模型构建

1) 特征变量筛选

在研究中,利用 GBRT 算法构建高山松地上生物量估测模型时,同样使用变量重要性评分方式对输入的 126 个特征变量的训练数据进行筛选。在进

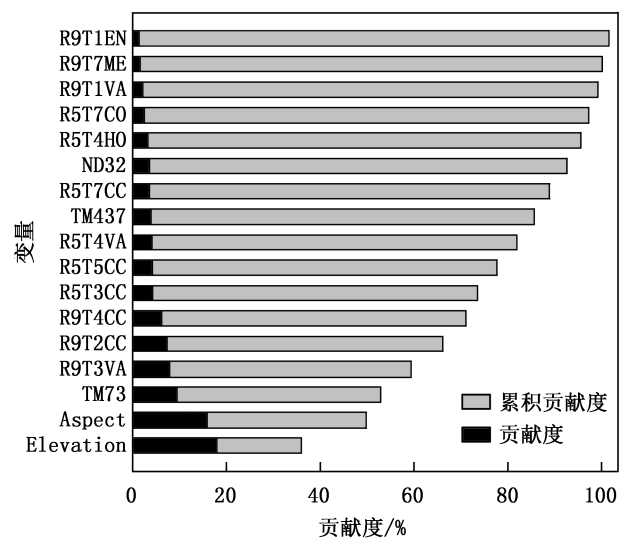


图 2 变量重要性指标

Fig. 2 Variable importance index

行第一次分析时,输入模型的变量为 126 个,当变量的累积贡献度达到 80% 时,共有 24 个特征参与重要性评分积累。使用前 24 个变量继续进行第二次模型构建,当变量的累积贡献度达到 80% 时,共有 20 个特征参与重要性评分积累。继续使用这前 10 个变量进行第三次模型构建,当变量的累积贡献度达到 80% 时,共有 7 个特征参与重要性评分积累。进一步使用这前 7 个变量继续进行第四次模型构建。

2) 模型拟合

可以看出,当模型进行第三次运算,参与运算的变量数为 10 时,模型具有最高的预测精度和最低的相对均方根误差。本研究将选择第三次运算结果进行高山松地上生物量估测模型的构建(表 3)。

表 3 梯度提升回归树算法模型多次构建效果

Tab. 3 Multiple construction effects of gradient boost regression tree algorithm model

运算次数	变量数	R^2	RMSE /($t \cdot \text{hm}^{-2}$)	P /%	rRMSE /%
第一次运算	126	0.98	13.99	62.12	33.65
第二次运算	24	0.92	28.33	71.65	32.70
第三次运算	10	0.99	9.88	70.07	26.69
第四次运算	7	0.86	36.39	67.26	46.07

此时参与模型构建的 10 个特征变量为:R5T5CC、TM73、R5T3CC、R9T2CC、Aspect、R5T7CC、Elevation、R9T1HO、R5T4ME、NDVI。在模型中变量的贡献度排序如图 3 所示。

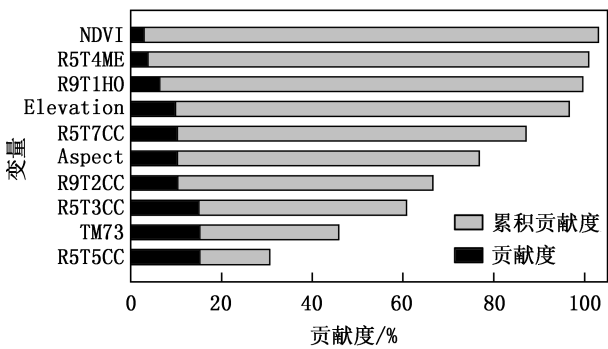


图 3 变量重要性指标
Fig. 3 Variable importance index

构成模型各特征变量中,R9 和 R5 分别代表灰度共生矩阵计算时大小为 9×9 和 5×5 的窗口;TM1、TM2、TM3、TM4、TM5、TM7 分别代表 Landsat5 TM 影像的第 1、2、3、4、5、7 波段,即 Blue 波段、Green 波段、Red 波段、NIR 波段、SWIR1 波段和 SWIR2 波段;CC 代表纹理特征相关性,HO 代表纹理特征均匀性,ME 代表纹理特征均值。

张加龙^[7]等人的研究表明,纹理因子对生物量的研究具有重要影响,其中 HO、CC 和近红外波段信息对生物量的影响最大。在非参数方建模法中,以 GBRT 算法作为高山松遥感生物量估测模型,其相对均方根误差、 P 值、 R^2 均优于 RF 和 KNN 算法。

3.2 高山松地上生物量估测模型比较分析

采用不同回归算法构建的估测模型对高山松地上生物量进行估测,估测效果见表 4。

表 4 基于不同回归算法构建的估测模型效果比较
Tab. 4 Effects of estimation models based on different regression algorithms

算法	R^2	RMSE/($t \cdot hm^{-2}$)	$P/\%$	rRMSE/%
GBRT	0.99	9.88	70.10	26.69
RF	0.89	32.26	66.10	40.92

从表 4 可以看出,通过分析两种算法的建模效果,GBRT 表现出了最优的预测精度(70.10%)和最小的相对均方根误差(26.69%),拟合优度达 0.99。因 GBRT 算法在对训练数据优秀的拟合度以及对预测数据的精准预测,本研究将选择 GBRT 算法对 5 期香格里拉市的高山松地上生物量进行反演和制图。

4 结果与分析

根据模型比对,选择 GBRT 算法评估香格里拉市高山松地上生物量。反演时提取遥感影像每个像

元对应的各个遥感因子值,根据最优模型计算得出每个像元的地上生物量值,使用高山松范围矢量数据进行裁剪,再进行统计并制图,最终得到各年份香格里拉市高山松地上生物量及其分布图(表 5,图 4)。

表 5 高山松地上生物量多年份反演结果
Tab. 5 Multiyear inversion results of aboveground biomass of *Pinus densata*

年份	分布面积 / hm^2	地上生物量 /($t \cdot hm^{-2}$)	地上生物量 总量/t
1987	171560.28	59.65	10232946.16
1992	171560.28	59.59	10223784.37
1997	170589.86	59.31	10117252.71
2002	170589.86	59.68	10180215.68
2007	174179.37	58.52	10193252.57

1987—1997 年间,香格里拉市高山松空间分布面积是净减少期间,同时,香格里拉市高山松分布区范围内的平均地上生物量呈减少趋势,二者的叠加影响直接导致此时段高山松地上生物量的净减少。80—90 年代,为了发展地方经济,盲目新建消耗森林资源的森工企业,加剧了森林资源的消耗。这时期香格里拉市高山松地上生物量总量由 1987 年的 1023.29 万 t 下降至 1992 年的 1022.38 万 t,1997 年下降至 1011.72 万 t,这 10 年间生物量净损失 11.57 万 t。1997 年以后,中国采取了十分严格的天然林保护制度,1998 年 8 月 5 日国务院关于退耕还林工作的紧急通知中明确规定,凡坡度在 25° 以上的坡耕地都要在 2003 年以前全部实行退耕还林。因此,1997—2002 年的 5 年间香格里拉市森林植被面积是净增长的,香格里拉市高山松地上生物量由 1997 年的 1011.73 万 t 增长到 2002 年的 1018.02 万 t,而这一系列森林保护措施成效也在 2007 年的高山松空间分布范围的大面积增长数据中体现出来。2002—2007 年的 5 年间,香格里拉市高山松空间分布面积由 2002 年的 $170589.86\text{ }hm^2$ 增长到 2007 年的 $174179.37\text{ }hm^2$,净增长 $3589.51\text{ }hm^2$ 。在这期间由于新苗增加,高山松平均单位面积地上生物量由 2002 年的 $59.68\text{ }t/hm^2$ 下降至 2007 年的 $58.52\text{ }t/hm^2$ 。尽管平均值减少,但因总面积的增长,香格里拉市高山松地上生物量由 1997 年的 1011.73 万 t 增长至 2007 年的 1019.33 万 t,净增长 7.6 万 t。尽管 1987—2007 年的 20 年间香格里拉市高山松地上生物量总量净增长了 7.60 万 t,但由于 1987—1997 年净损失 11.57 万 t,因此,研究区高山松地上生物量仍然未恢复到 1987 年的水平。

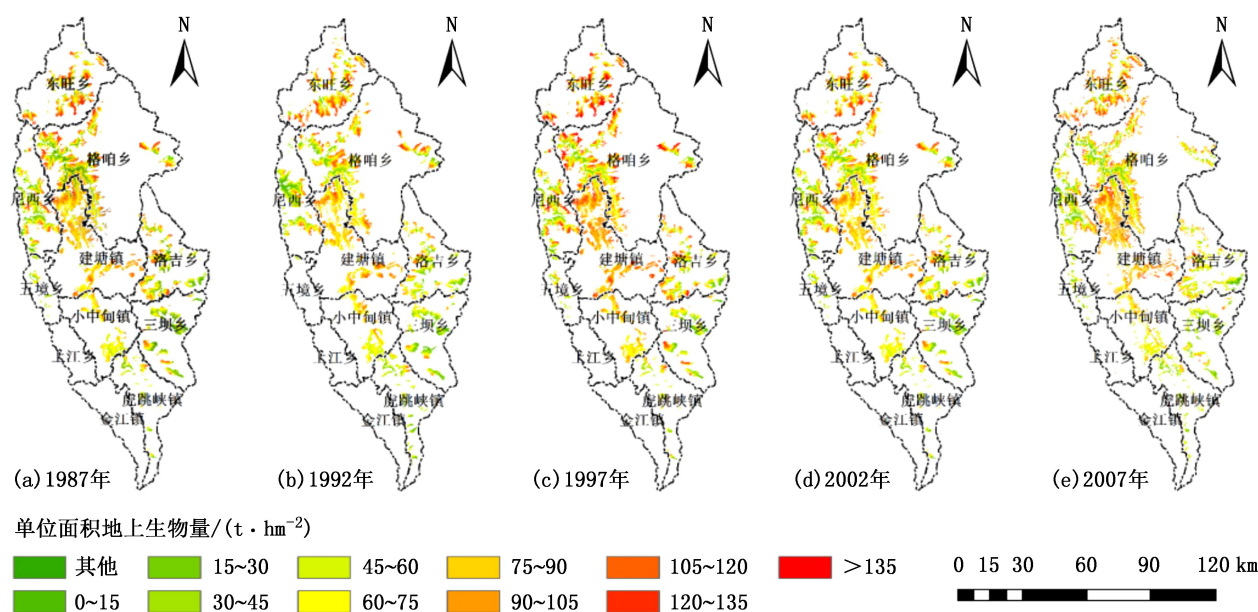


图 4 香格里拉市各年份高山松地上生物量分布

Fig. 4 Distribution of aboveground biomass of *Pinus densata* in different years in Shangri-La City

5 结论与讨论

基于 GBRT 算法构建的高山松地上生物量估测模型在拟合效果方面表现出较好的拟合度,同时其在预测效果方面也表现出最好效果。本研究中参与模型训练的样地实测数据相比于整个香格里拉市较为广阔的高山松分布区域仍显得不足(本文获得的样地记录数据为 96 条)。越多的数据训练意味着越精确的模型构建,更能精确反映出香格里拉市各年份高山松地上生物量分布。

本研究中存在诸多不足,对后续研究提出如下建议:

1) 可以考虑采用技术手段消除不同时间传感器成像时受到传感器状态、大气状态、地表状态的影响,使影像之间的差异得到消除,得到更能够反应植被与植被光谱特征之间的更真实关系,不仅有利于提高数据质量,还能将森林资源调查成果的历史存档数据有效利用起来,充实森林地上生物量的样地可用数据。

2) 充分利用现有研究成果,将更多的提高数据质量的方法引入到森林地上生物量估测研究中来,进一步提高森林地上生物量遥感估测精度。

3) 进一步加强林业调查资料的搜集与整理,并进行电子归档,逐步提高研究深度。

参考文献:

[1] 李明泽,于欣彤,高元科,等. 基于 SAR 极化分解与 Landsat 数据的森林生物量遥感估测[J]. 北京林业大

学学报,2018,40(2):1-10.

[2] 温小荣. 森林资源二类调查关键技术与方法的研究[D]. 南京:南京林业大学,2017.

[3] 温小荣,孟雪,刘俊,等. 基于不同立地质量的杉木生物量遥感估测[J]. 林业科学研究,2016,29(4):494-499.

[4] 陆驰. 基于 Landsat 的香格里拉市高山松地上生物量及其动态变化建模研究[D]. 昆明:西南林业大学,2017.

[5] 孙雪莲. 基于 Landsat8-OLI 的香格里拉高山松林生物量遥感估测模型研究[D]. 昆明:西南林业大学,2016.

[6] 徐云栋,舒清态,李圣娇,等. 基于优化 BP 神经网络的香格里拉高山松蓄积量模型研究[J]. 西北林学院学报,2015,30(6):190-195.

[7] 张加龙,胥辉,陆驰. 应用 Landsat8 OLI 和 GBRT 对高山松地上生物量的估测[J]. 东北林业大学学报,2018,46(8):25-30.

[8] 陆驰,张加龙,王爱芸,等. 基于 Landsat TM 的香格里拉市高山松生物量估测重建[J]. 林业调查规划,2016,41(6):1-7.

[9] 张鹏,王学强. 大气辐射校正软件 LEDAPS 与 FLAASH 对比研究[J]. 安徽农业科学,2014,42(10):3105-3108.

[10] ZHANG J L, LU C, XU H, et al. Estimating aboveground biomass of *Pinus densata*-dominated forests using Landsat time series and permanent sample plot data[J]. Journal of Forestry Research, 2019, 30(5):1689-1706.

[11] 张加龙,胥辉. 基于遥感的高山松连清固定样地地上生物量估测模型构建[J]. 北京林业大学学报,2020,42(7):1-11.

责任编辑:许易琦